A Probabilistic Formulation for Hausdorff Matching

Clark F. ()]S011
Jet Propulsion Laboratory, California Institute of Technology
Mail Stop 107-102, 4800 Oak Grove Drive, Pasadena, CA 91109
http://robotics .jpl.nasa.gov/people/olson/homepage .html

Abstract

Matching images based on a Hausdorff measure has become popular for computer vision applications. However, no probabilistic model has been used in these applications, This limits the formal treatment of several issues, such as feature uncertainties and prior knowledge. In this paper, we develop a probabilistic formulation jor *Hausdorff matching in terms oj maximum likelihood estimation. This formulation yields several benefits with respect to previous Hausdorff matching formulation. The techniques are applied to a mobile robot self-localization problem,*

1 Introduction

The **use** of **variants** of the Hausdorff distance has recently become popular in image matching applications (see, for example, [4, 6,8, 13, 14, 15]). J$'bile these methods have been largely successful, they have lacked a probabilistic formulation of the matching process, and this has made it difficult to incorporate probabilistic information, such as feature uncertainties and the prior probability distribution of model positions, into these applications. This work addresses these issues by introducing a probabilistic formulation of Hausdorff matching.

After a brief review of Hausdorff matching techniques, we describe a probabilistic formulation of Hausdorff matching based on the principal of maximum likelihood estimation. In this formulation, we seek local maxima of the likelihood function over the possible model positions, assuming that the model appears in the image. Note that this formulation can be applied even when the model does not appear in the image or appears multiple times. We must simply set the criterion determining which model positions are reported as likely hypotheses appropriately, When a particular probability distribution function (PDF) is introduced for the distance of each model feature from an image feature in the image, this formulation yields the conventional Hausdorff matching method. Alternate PDF's yield new and interesting variations of the method.

This probabilistic formulation of Hausdorff matching yields several benefits, It allows the incorporation of prior knowledge, such as the prior probability distribution of model positions, into the matching process. It also allows formal treatment of feature uncertainties in the search for likely model positions. In addition, with this formulation we can consider arbitrary probability distribut ions for the locations of the image features, rather than the simple two-valued support function that corresponds to conventional Hausdorff matching methods.

We discuss efficient techniques for searching the pose space in this formulation and give experimental evidence that indicates improved accuracy in the recognition and localization of objects in images is achieved. Finally, we apply these techniques to a mobile robot self-localization application that performs matching between terrain occupancy maps to determine the robot's position.

2 Hausdorff matching

This section reviews a variation of the Hausdorff distance commonly used to perform image matching, as well as the application of this measure to matching in binary images and an efficient search strategy for finding the relative image positions where the measure meets some criterion.

2.1 Hausdorff measure

For two sets of points A an,] B, the directed Hausdorff distance from A to B is:

$$h(A,B) = \max_{a\in A}\min_{b\in B} ||a-b||, \qquad (1)$$

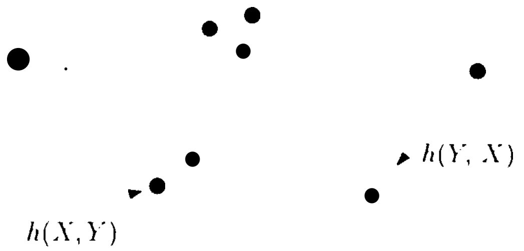


Figure 1: Directed Hausdorff distance between point sets. The white points are set X and the black points are set Y.

where $||\cdot||$ is any norm. This yields the maximum distance from a point iu set A to its nearest point in set $B$and is illustrated in Figure 1. Notice, however, that a single outlier in A can change this distance by an arbitrary amount. For image matching, where A is usually a set of model points and B is a set of image points, we wish to allow outliers. It is thus common to use the partial distance [3]:

$$h_K(A, B) = \underset{a\in A}{K^{\text{th}}} \min_{b\in B} ||a -- b|| \qquad (2)$$

This yields the Hausdorff distance among the K points in A that best match points iu B (and thus allows $|A|$ - K outliers iu the set .4).

A variation on the partial Hausdorff distance is to determine the maximum number of points in the model such that distance is below a given error threshold:

$$h_K(A, B) \leq \delta \qquad (3)$$

Let $K_\delta(A, B)$ denote the maximum K for which (3) is true. The ratio $F_\delta(A, B) = \frac{K_\delta(A,B)}{|A|}$ is called the *Hausdorff fraction*, since it is the fraction of the points in A that match a point in B up to the error δ. This formulation is easy to work with, siuce $K_\delta(A, B)$ is simple to compute, and we examine this variation of Hausdorff matching in this paper,

Note that pre-setting some maximum error δ, and determining the model positions such that F_δ is shove some threshold T, yields equivalent results to setting the model fraction to T aud determining the model positions with partial Hausdorff distance no greater than δ. This formulation of the Hausdorff metric does not change the solutions that are found.

2.2 Application to binary images

We concentrate on the application of these techniques to binary digital images (e.g. image edge maps). Each pixel in such an image takes a value of O or 1. We say that the pixels with a value of 1 are *occupied* aud those with a value of O are *unoccupied*,

Let M he a model image or template and I be an image that may contain au instance of the model. Both M and I can be considered to be discrete sets of points corresponding to the locations of the occupied pixels in the image or template. Let t **be a** particular position of the model with respect to the image. This model position can be thought ofa function that maps tile model points into the image; $t(M)$ is thus the set of model points after mapping them according to t.

Now, consider the dilation of the image by the structuring element S_δ that consists of all of the pixels within δ of the origin with respect to some norm. The dilated image, $I_\delta = I \oplus S_\delta$ (where $\oplus$ denotes the Minkowski sum or morphological dilation operator), has an occupied pixel at each location that is within δ of an occupied pixel in the original image. Let $I_\delta(m)$ denote tile value of I_δ (i.e. 0 or 1) at the position of some model pixel, m. We can write the Hausdorff fraction (as a function of the model position) as follow's:

$$F_\delta(t(M), I) = \frac{1}{|M|} \sum_{m\in t(M)} I_\delta(m) \qquad (4)$$

2.3 Efficient search strategy

The best known search strategy for locating model positions that satisfy some criterion with respect to the Hausdorff fraction is a multi-resolution search strategy that examines a hierarchical cell decomposition of the space of possible model positions [12]. This method divides the space of model positions iuto rectilinear cells and determines which cells may contain a position satisfying the criterion using some test. The cells that pass tile test are divided iuto subcells, which are examined recursively. The rest are pruned.

The key to this method of searching the parameter space is a quick method to conservatively test whether a cell can contain a position satisfying the criterion. This test is allowed to fail to rule out a cell that does not contain any positions satisfying the criterion, but it should never rule out a cell that does contain such a position.

It is typical in this method to consider only the model positions in some underlying discretization of the pose space. When a cell of this space is reached that contains a single position in the discretization, this position is tested explicitly.

In order to develop an efficient testing mechanism for determining whether a cell can be pruned, it is useful to consider the distance transform of the image. The distance transform of a binary image measures the distance from each pixel in the image to the closest occupied pixel [1 1]. Denote the distance transform of image $I(x, y)$ by $D_I(x, y)$.

To test a cell C of possible model positions, the discrete pose c closest to the center of the cell is first determined. The maximum distance between the location to which a model pixel is mapped into the image by c and by any other pose in the cell is then computed. We call this distance the *image-mapped radius* of the cell and denote it A_c:

$$\Delta_C = \max_{p \in C} \max_{m \in M} ||p(m) - c(m)||$$

Now, if we seek positions at which $K_\delta(t(M), I)$ is no less than T, then, to test the cell, we count the number of model points for which the probe into the distance transform at the appropriate location is no larger than δ + AC. If this number is less than T, then we can prune the cell, since it cannot contain a model position that matches T pixels in the model to pixels in the image up to the error δ [12].

When a cell cannot be pruned, it is divided into multiple subcells, and the procedure is applied recursively to each of the subcells. This process continues until all of the cells in the pose space have been exhausted.

3 Probabilistic formulation

We now describe a probabilistic formulation of Hausdorff matching using the principal of maximum likelihood estimation. To formalize the problem, let us say that we have a set of model features, $M = \{\mu_1, \ldots, \mu_m\}$ and a set of image features, $I = \{\nu_1, \ldots. \nu_n\}$. Let $t \in T$, be a random variable describing the position of the model in the image. This makes an implicit assumption that exactly one instance of the model appears in the image. However, we shall see that the cases where the model does not appear, or the model appears in multiple instances, can be easily handled in this formulation.

To formulate the problem in terms of maximum likelihood estimation of the model position, we must have some set of measurements that are a function of the position of the model. We use the distance from each model pixel (at the position specified by t) to its closest occupied pixel in the image as our set of measurements. Each distance can be found simply by looking up the position of the model pixel in the distance transform of the image. These distances are random variables that we denote by $D_1, \ldots. D_n$. Jf'bile these distances arc not independent, we model them as such. Recent work on determining the probability of a false positive for Hausdorff matching [1, 10] provides support for treating the model features independently. With this approximation, we can formulate the likelihood function for t as the product of the probability distributions of these distances:

$$L(t) = \prod_{i=1}^{m} p(D_i; t), \tag{5}$$

where $p(D_i; t)$ is the probability distribution function (PDF) of D_i as a function Oft. Taking the logarithm of (5) yields:

$$\ln L(t) = \sum_{i=1}^{m} \ln p(D_i; t) \tag{6}$$

With a particular PDF this yields a measure equivalent to $K_\delta(t(M), I)$. The PDF necessary for this to be the case satisfies:

$$\ln p(D_i; t) = \begin{cases} k_1 + k_2 & \text{if } D_i \le \delta \\ k_1 & \text{otherwise} \end{cases} \tag{7}$$

This probability distribution function is two-valued as in the conventional Hausdorff matching formulation. If there is support for the model feature in the image at this position (i.e. an image feature lies with δ of it), then some constant probability is assigned to $p(D_i; t)$, otherwise some smaller constant probability is assigned to $p(D_i; t)$. The precise values of k_1 and k_2 are unimportant in this analysis as long as $k_2 > 0$. In practice, we use k_1 = O and k_2 = 1.

Now, let us address the assumption implicit in this formulation that the model appears exactly once in the image. If we have are seeking models that may appear more than once in an image, or not at all, we must only set some threshold on (6), as is usually done in Hausdorff matching formulations. The model positions that surpass the threshold correspond to the likely positions of the model in tile image.

4 Using the probabilistic formulation

This section explores some of the advantages that arc yielded by the probabilistic formulation of Hausdorff matching.

4.1 Uncertainty in the image features

The probabilistic formulation of Hausdorff matching allows the formal treatment of uncertainties in the image features. For example, we may have a feature detector that yields uncertainty estimates for the position and/or the existence of the feature. A feature that is less likely to exist in the image, or for which the position estimate is inaccurate, might be weighted less in the matching process, and for features with inaccurate position estimates, the allowable positions of model points that can match them should be larger.

In order to treat these issues, we modify the PDF of our measurements, D_i. Each image point (indexed by j) may be assigned a probability of existence, e_j, and a uncertainty radius, r_j. We now define:

$$\ln p(D_i;t) = \max_{1 \le j \le n} \begin{cases} k_1 + k_2 \frac{e_j}{{r_j}^2} & \text{if } \|t(m_i) - i_j\| \le r_j \\ k_1 & \text{otherwise} \end{cases} \tag{8}$$

This allows image features to be weighted by their probability of existence and to contribute to an arbitrary radius in the image. While image features with larger positional uncertainties contribute to a larger area in the image, their contribution at each position is less. Once again, the constants, k_1 and k_2, are irrelevant to finding local maxima and we use k_1 = O and $k_2 = 1$.

We can vary the probability distributions more significantly, if desired. For example, we may use a normal distribution with a constant additive term:

$$p(D_i;t) = \max_{1 \le j \le n} k_1 + k_2 e^{-\|t(m_i) - i_j\|^2 / k_3} \tag{9}$$

This distribution models the case where the error in feature localization has approximately a Gaussian distribution. The added constant allows for cases when the feature is not found at all.

Note that the efficient search strategy discussed above does not work directly with these changes. We require some modification to the search strategy to perform matching with this formulation.

4.2 Prior probabilities of model positions

In some applications, we have prior knowledge of the likelihood of various model positions being correct. For example, in tracking applications (e.g. [2, 5]) we may use the previous position of the object being tracked and its velocity to predict the next position of the object. It is often reasonable to model the probability distribution function of the error in position with a normal distribution:

$$E(\delta_x, \delta_y) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{\delta_x^2 + \delta_y^2}{2\sigma^2}}, \tag{1o}$$

where δ_x and δ_y are the differences in the position from the position estimate, and σ is the standard deviation of distribution.

In tile case where the prior probability of each model position is not uniform, let $f(t)$ be the prior probability of position t. We now have:

$$L(t) = f(t) \prod_{i=1}^{m} p(D_i;t) \tag{11}$$

Taking the logarithm yields:

$$\ln L(t) = \ln f(t) + \sum_{i=1}^{m} \ln p(D_i;t) \tag{12}$$

It is relatively easy to incorporate this information into the efficient search strategy. When we compute Δ_C, we must only include the maximum decrease (or minimum increase) in the first term 011 the right side of (12) when determining whether a cell can be pruned to ensure that we do not rule out any cell than can contain a valid position.

The use of prior information as to the likelihood of various model positions being correct yields the additional benefit that we have bounds in the image on the space we need to search. We need not examine any position for which in $f(t)$ is so small that the sum with the best possible score for each of tile model pixels could not surpass the score for the best known position (or some threshold if we seek all positions with scores above the threshold).

5 Efficient algorithm

As noted above, some of the possible modifications that can be made to the matching formulation require that the search strategy be rethought. This section discusses techniques that allow the space of possible model positions to be searched efficiently for positions that satisfy some matching criterion according to the probabilistic formulation of Hausdorff matching.

Let us first note that a brute force method can be constructed by determining, for each pixel location in the image, the value of $\ln p(D_i;t)$, since $p(D_i;t)$ is independent of the particular model feature; only the position that t maps m_i into the image is important. We can thus compute a transform of the image, denoted by $P(X) = \{p(X;t) \mid t(m) = X\}$, where $Y = [x\, y]^T$ is a pixel location in the image, according to Equation (8) or Equation (9). We call this the *feature probability transform* of the image. Each possible position of the model can then be tested by probing this transform at the location that the position maps each model feature, summing them, and determining if the sum meets the criterion.

Now, to search the space efficiently, we adapt the multi-resolution search strategy discussed previously, where we attempt to prune large cells of the transformation space. Recall that in this search strategy, we compute, for each cell that is examined, the discrete model position closest to the center of the cell and image-mapped radius of each cell, denoted by c and A$_c$, respectively. l'hen, each of the model features is tested todetermine if there could be a position within the cell where the modelfeature is matched byan image feature up to the allowable error.

In the new formulation, we instead want to determine the maximum conditional probability that a model feature could have with respect to any model position in the cell. While it is not efficient to compute these values upon demand, they are a function only of the image-mapped radius of the cell and the position in the image. If we take care to ensure that all of the cells at each level of the search have the same dimensions, we can efficiently compute all of the values at once.

Let Δ_L be the maximum image-mapped radius over the unpruned cells at level L.Note that for many transformation spaces (translations, for example) AC depends only on the size of the cell, not the cell position. So, if all of the cells at level 1, have the same dimensions, then they also have the same image-mapped radius. We compute, for each level of the tree, a dilation of $P(X)$ that yields, for each pixel, the maximum value over the prescribed distance, AI,:

$$P_{\Delta_L}(X) = \max_{Y \in \{X\} \oplus S_{\Delta_L}} \mathrm{P}(\,\}\,) \qquad (13)$$

Now, if we sum the probes of $P_{\Delta_L}(X)$ at the locations where c maps each of the model points and the result still does not satisfy the matching criterion, then we can prune the entire cell. We must precompute each relevant $P_{\Delta_L}(X)$ prior to the search, if a deptll-first or best-first search is used, but we need only store a single $P_{\Delta_L}(X)$ at a time, if a breadth-first search strategy is used instead. The remainder of the search strategy remains the same.

6 Results

This section discusses the results of applying these techniques to both a synthetic problem, where we are concerned with matching two-dimensional data, and a real application, where we localize a mobile robot by matching tllree-dimensional range maps.

6.1 Synthetic experiments

We first tested these techniques in controlled experiments where exact ground truth was available, since the image feature data was generated synthetically. We chose a simple problem domain (translation of isolated feature points) under demanding conditions to demonstrate the superiority of the probabilistic formulation. This experiment generated random model features (to subpixel accuracy). The model was translated randomly and placed in the image with considerable occlusion, clutter, and noise. Scc [9] for details.

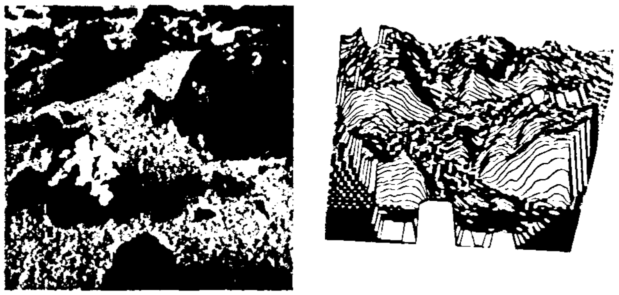
Over 10000 trials, the conventional Hausdorff matching method yielded 1293 instances where an incorrect match had a higher score than the correct match, while the probabilistic formulation, using **a** probability distribution similar to (9), yielded 71.5 such failures on the same images. The probabilistic formulation thus yielded superior recognition of the feature patterns.

We also tested the localization accuracy of the techniques. Note that a lower bound on the average accuracy of matching of 0.25 pixels in each direction exists, since matching is performed only to pixel accuracy. In the successful trials, the probabilistic formulation yielded an average localization error of 0.36 pixels in each direction, while the conventional method yielded an average error of ***0.48*** pixels. The average error of the conventional method was thus over twice as far from the theoretical minimum as with the probabilistic formulation.

6.2 **Mobile robot localization**

While the synthetic problem described above yields positive data with respect to the performance of the probabilistic formulation of Hausdorff matching, the real test, of course, is in real applications. We have previously implemented a mobile robot localization method using conventional Hausdorff matching methods [8]. Here we compare this system to a new implementation using the probabilistic formulation.

The motivation for studying this problem is to allow the next generation of Mars rovers to have greater autonomy from the lander and from human operators. The basic method that is used is to generate a range map of the terrain near the robot through stereo vision [7]. This range map is transformed into a three-dimensional occupancy map describing the terrain (see Figure 2) and it is then compared against a previously



(a) (1))

Figure 2: Range maps are computed using stereo vision. (a) Left image of a stereo pair. (b) Surface extracted from the stereo pair.

generated occupancy map of the terrain to determine the relative position between the maps. For example, it can be compared to a range map generated from previous robot positions, or to a map generated prior to the robot activity by some other means [8]. While the matching techniques described here have been discussed in terms of two-dimensional edge maps, the generalization to three-dimensional surface maps is straightforward.

In an experiment over 13 camera positions, where the ground truth was measured by hand, the previous implementation using the conventional Hausdorff matching method had an average error of 0.050 meters, while the new implementation yielded an average error of 0.042 meters. It is likely that human error in collecting the ground truth is responsible for a significant amount of the remaining error. In similar experiments where the cameras were panned by 25 degrees, but were not translated, the error was reduced from 0.011 meters to 0.004 meter-s. The probabilistic formulation of Hausdorff matching thus yielded significantly improved results in this problem domain.

7 Summary

The primary contribution of this paper is a oew formulation of Hausdorff matching in terms of maximum likelihood estimation. This formulation seeks local maxima in the likelihood function of posit ion of the model with respect to the image, where it is implicitly assumed that the n iodel appears in the image. However, this formulation can be applied equally well when the model does not appear in the image if an appropriate threshold is used to determine which locations are output as likely model positions.

This formulation yields several advantages over previous work in this area. First, feature uncertainties, in both the posit ion and existence of the features, cau be t rest ed formally in the framework. Second, smoothly varying probability dist ribution functions can be used that eliminate the sharp boundary inherent in the conventional two-valued support function. In addition, it is simple to incorporate prior knowledge shout the probability distribution of model posit ions in the matching process in this formulation.

We have described new techniques for performing matching efficiently in this formulation. Experiments on synthetic data imply that the new techniques yield performance superior to the standard formulation with respect to both recognition and localization. Finally, we have applied this technique to the self-local ization of a mobile robot in a natural environment using range maps from stereo vision. Improved results were also obtained in this domain.

Acknowledgments

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

[1] W. E. I,. Grimson and I). P. Huttenlocher. Analyzing the probability of a false alarm for the Hausdorff distance under translation. In *Proceedings of the Workshop on Performance versus Methodology in Comp uter Vision*, pages 199 20s, 1994.

[2] I). P. Huttenlocher. Using two-dimensional models to interact with the three-dimensional world. In Proceedings *of the International NSF-ARPA Workshop* on *Object Representation in Comput er Vision*, pages 109 124, 1995.

[3] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the Hausdorff distance. *IEEE Transactions* on *Pattern Analysis and Machine Intelligence*, 15(9):850- S63, September 1993.

[4] 1). P. Huttenlocher, M. F,. Leventon, and W. J. Rucklidge. Visually-guided navigation by comparing two-dimensional edge images. In *I'rote.ed!ngs of the IEEE Conference on ComputerVision and Pattern Recognition*, *pages* 8-12-8-17, 1994.

[5] 1). P. Huttenlocher, J. J. Noh, and W. J. Rucklidge. Tracking non-rigid objects in complex scenes. In Proceedings *of the International Conference on Computer Vision*, pages 93 101, 1993.

[G] K. Kedem and Y. Yarmovski. Curve based stereo matching using the minimum Hausdorff distance. In *Proceedings* Of *the ACM Symposium on Computational Geometry, pages 15-18,* 1996.

[7] L. Mattl ties. Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation. *International Journal of Computer Vision,* 8(1):71-91, July 1992.

[8] C. F. Olson. Mobile robot self-localization by iconic matching of range maps. In Proceedings *of the International Conference on Advanced Robotics,* 1997.

[9] C. F. Olson. A probabilistic formulation for Hausdorff matching. Technical *report,* Jet *Propulsion* Laboratory, *1997.*

[10] C.F. Olson and D. P. Huttenlocher. Automatic target recognition by matching oriented edge pixels. *IEEE Transactions on Image Processing,* 6(1):103--113, January 1997.

[11] A. Rosenfeld and J. Pfaltz. Sequential operations in digital picture processing. *Journal of the ACM,*13:471--494, 1'366.

[12] W. J. Rucklidge. *Efficient Visual Recognition Using the Hausdorff* distance. Springer-Verlag,1996.

[13] E. Saber and A. M. Tekalp. Image query-by-exa[[,ple using region-based shape matching. In *Image and Video Processing IV, Proc. SPIE* 2666, pages 200-211, 1996.

[14] X. Yi and O. I. Camps. Robust occluding contour detection using the Hausdorff distance. In Proceedings *of the IEEE Conference on Computer Vision and Pattern Recognition,* pages '362-968, 1997.

[15] R. Zabih, J. Miller, and K. Mai. Video browsing using edges and motion. In Proceedings *of the IEEE Co nfe rence on Compu ter Vision and Pattern* Recognition, pages 439 446, 1996.